《心理学报》论文自检报告

请作者填写以下内容，粘贴在稿件的首页。

1. 请以“研究亮点”的形式列出最多三条本研究的创新性贡献，总共不超过 200 字。
《心理学报》的目标是发表“既科学优秀，又具有广泛兴趣和意义”(be both scientifically excellent and of particularly broad interest and significance)的心理学前沿研究。如果您的研究只有小修小补的贡献，没有尝试开创新的研究领域(new areas of inquiry)或提出独到见解和创新视角(unique and innovative perspectives)，特别纯粹只是研究没有明确心理学问题的算法或技术的工作，这类研究被本刊接受的机会小。建议另投它刊。
答： 1.采用内隐启动范式，选择英语水平低的汉语母语者，排除了二语的音韵编码单元可能对汉语口语词汇产生过程的影响。2.在汉语口语词汇产生过程中同时探测到了音节与音素效应，说明讲话者在词条选择之后，首先提取音节信息，再将音节分解为音素。3. 音节效应发生在音韵编码阶段，而音素效应可能发生在音韵编码阶段后期或语音编码阶段。

2. 作者已经投稿或发表的文章中是否采用了与本研究相同的数据？如果是，请把文章附上审查。(我们不赞成作者用同一数据发表多篇变量相同的文章，也不赞成将一系列的相关研究拆成多个研究来发表的做法。)
答：否

3. 管理、临床、人格和社会等领域仅有自我报告(问卷法)的研究，需要检查数据是否存在共同方法偏差(common method bias)。为控制或证明这种偏差不会影响研究结论的效度，你使用了什么方法？采取了哪些措施？(共同方法偏差的有关文献可参见：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract894.shtml)基于横断数据，仅有自我报告，仅仅在方便样本中施测，这样的研究数据易取得，但通常创新性价值不大，被本刊接受的机会小。
答：本研究是实验研究，不存在这一问题。

4. 以治疗疾病为目的临床实验，建议在收集数据前预先备案(pre-register)。也鼓励其他实验研究预先备案。备案要求写出所有的研究假设及其支持，以及实验/干预的详细过程和步骤。本期刊的预先备案网站 https://osf.io/。如果您的研究有预先备案，会显著增加被录用的机会。如果您的研究有预先备案，请写出登记号。预先备案的重要性可参考 https://osf.io/5awp4/ 。
答：本研究是对健康人口语产生过程的研究，不是以治疗疾病为目的的临床实验。

5. 是否报告并分析了效果量(effect sizes；如：t 检验：Cohen’s d；方差分析：Eta 平方或偏 Eta 平方；标准化回归系数)？(很多研究只是机械地报告了效果量，但没有做必要的分析或说明，如效果量是大中小？有什么理论意义或应用意义？)。(在 google 中搜索“effect size calculator”，可搜到许多计算方便的 APP。效应量的有关解释，中文可参考：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract1150.shtml；英文可参看：http://www.uccs.edu/lbecker/effect-size.html
是否报告统计分析的 95%CI？(如，差异的 95%CI；相关/回归系数的 95%CI)置信区间的有关计算和绘图可参考 https://thenewstatistics.com/itns/esci/)
答：研究中涉及的方差分析结果全部报告了相应效果量(偏 Eta 方)。同样报告了 95% CI，论文结果部分不同时间窗口的音节与音素效应(图 3)的误差棒为 95%CI。

6. 请写出计划的样本量，实际的样本量。如果二者有差别，请写出理由。以往心理学研究中普遍存在样本量不足导致的低统计功效(power)问题，我们建议在论文的方法部分解释您计算及认定样本量的依据。应该以有一定依据的效果量(effect size)、期望的功效来确定样本量，并报告计算用软件或程序。样本量计划的理由和做法可参考 https://osf.io/5awp4/
答：计划样本量 20 名被试，实际样本量 20 名被试。

7. 假设检验中，如果是零假设显著性检验(NHST)，需报告精确 p 值而不是 p 的区间(小于 0.001 的报告区间，其他报告精确 p 值)。你的论文是否符合该项要求？如果是贝叶斯因素，是否已报告其对先验分布假定的敏感性？
答：本论文符合要求。

8. 为保证论文中数据报告的完备性，统计分析中如果剔除了部分数据，是否在文中报告？原因是什么？包含这部分数据统计结果如何变化？统计分析中是如何处理缺失数据的？使用量表时是否删除了其中的个别题目？原因是什么？如果包含这部分题目统计结果会如何变化？是否有测量的项目或者变量没有报告，原因是什么？请写出在论文中的位置。
答：统计分析中涉及的所有部分数据的删除在论文的结果部分第一段、第二段和第三段均有详细介绍，除文章中说明的删除的数据外，没有额外的数据删除，文章涉及的所有数据中

没有缺失数据的情况。

9. 研究用到的未经过同行评议和审查的实验材料、量表或问卷，是否附在文件的末尾以供审查？如果没有，请写出理由。如果该文发表，您是否愿意公开此程序与其他研究者共享？
答：本研究中所使用的所有实验材料均为经过同行评议和审查的材料，具体材料来源详见方法中材料部分。

10. 您的研究的原始数据是否可以上传备审稿人和编辑核查用？如果不能，请写出理由。如果该文发表，您是否愿意公开数据供其他研究者共享？
答：本研究原始数据可上传以备核查，同时愿意公开数据供其他研究者共享。

11. 您的研究如果用到了人类或动物被试，是否得到所在单位伦理委员会的批准？如果是，请把扫描版发至编辑部邮箱。如果否，请说明理由。
答：是。

12. 是否依据编辑部网站发布的“英文摘要写作注意事项”撰写 400~500 个单词的英文大摘要？英文题目和摘要是否已请英语好的专业人士把关或者已送专业 SCI/SSCI 论文编辑公司修改润色？
答：是。

汉语口语词汇产生的音韵编码单元：内隐启动范式的ERP研究[1]

王雪娇 张清芳

(中国人民大学心理学系, 北京 100872)

摘 要 本研究中选择英语水平低的汉语母语者，排除作为二语的英语音韵编码单元(音素)可能对汉语口语词汇产生过程的影响后，运用事件相关电位技术，考察了汉语口语产生过程中音节和音素效应的时间进程。实验采用内隐启动范式，要求被试看到提示词之后说出对应的目标词。事件相关电位分析结果发现，音节效应出现在线索词呈现后的100~400ms之间，音素效应出现在500~600ms之间，波形表现为相关条件比无关条件的波形更正。这表明在词汇选择之后的音韵编码阶段讲话者首先提取的单元是音节，而音素效应出现的时间窗口可能是音韵编码阶段后期或者是语音编码阶段，结果验证了合适编码单元假说的观点。

关键词 口语词汇产生；音韵编码；合适编码单元假说；音节；音素

1 引言

言语产生是指将概念或思想转换为语言进行输出的过程，包括了三个阶段：第一是概念化过程(Conceptualization)，说话者明确要用言语表达的概念是什么；第二是言语组织阶段(Formulation)，讲话者根据概念选择恰当的词汇，建立恰当的语法结构和发音结构；第三是发音阶段(Articulation)，选择的词汇通过一定的肌肉运动程序用声音表达出来(Dell, 1986; Roelofs, 1997; Levelt, Roelofs, & Meyer, 1999)。言语组织是口语产生中的核心阶段，也被称之为词汇通达，包括了两个阶段：词条选择(Lemma Selection)和单词形式编码(Word-form Encoding)。单词形式编码过程可进一步细分为词素音位编码(Morphophonological Encoding)、音韵编码(Phonological Encoding)和语音编码(Phonetic Encoding)三个过程。在词素音位编码过程，讲话者根据词条选择阶段所得到的词汇选择相应的词素及其句法特征。在音韵编码过程，根据词素选择音段和节律结构，并进行音节化(Syllabification)，将音段与节律结构中的音节节点联系起来。在语音编码过程选择音节程序节点为发音做好准备(Dell, 1986; Roelofs, 1997; Levelt et al., 1999)。

音韵编码的加工单元一直是言语产生研究的争论焦点之一(印欧语系：Dell, 1986;

[1] 收稿日期：2019-03-29
北京市社会科学基金重点项目(16YYA006)和中国人民大学科学研究基金项目(中央高校基本科研业务费专项) (18XNLG28)项目资助。
通信作者：张清芳, E-mail: qingfang.zhang@ruc.edu.cn

Levelt et al., 1999; Meyer, 1990, 1991; Forster & Davis, 1991; Kinoshita & Woollams, 2002; Malouf & Kinoshita, 2007; Schiller, 2008; Damian & Dumay, 2007, 2009; Damian & Bowers, 2003; Jacobs & Dell, 2014; 汉语：J.-Y. Chen, Lin, & Ferrand, 2003; O′Seaghdha, J.-Y. Chen, & T.-M. Chen, 2010; J.-Y. Chen, O'Seaghdha, & T.-M. Chen, 2016; Qu, Damian, & Kazanina, 2012; J. Wang, Wong, S. Wang, & H.-C. Chen, 2017; Zhang & Damian, 2019) 。口语产生的两大理论模型对于音韵编码的加工单元提出了不同的观点。根据语误分析的结果，Dell (1986)提出的单词形式编码模型认为：音韵编码的单元包括了音素(phoneme)[2]、音节(syllable)[3]、音节的组成成分及其语音特征。在音韵编码过程中音节的各个组成部分(首音、核心元音和尾音)同时得到激活后被插入音节框架结构。当一个词素包含多个音节时，音节的加工是从左至右序列进行的。Levelt 等(1999)的模型认为音韵编码的单元是音素，单个音节内部的音韵编码是一个增长式的编码过程，从音节的首音开始到核心元音最后到尾音。Dell（1986）认为音韵编码的单元可以是音节或音素，而 Levelt 等（1999）认为音韵编码的单元只是音素。两种观点的不同在于，前者认为音节内部的音素可以同时被激活并插入音节框架，而后者认为音素以序列的方式从左至右被插入音节框架。

在印欧语言的口语产生中，大多数研究支持音素是音韵编码的加工单元。语误分析和以反应时作为指标的各种任务(包括掩蔽启动范式、内隐启动范式以及图画-词汇干扰实验范式等)的研究结果都证实了这一观点。研究发现被试的语误主要表现为音素的遗漏、替换与置换，而极少发生音节的误用(Dell, 1986; Shattuck-Hufnagel, 1979)。采用内隐启动范式，研究者发现当连续说出的目标词的首音素相同时，其反应时短于首音素不同时，出现了首音素促进效应(Alario, Perre, Castel, & Ziegler, 2007; Damian & Bowers, 2003; Jacobs & Dell, 2014; Meyer, 1991)。在形容词-名词命名任务中，当形容词的首音素与名词的首音素相同时(例如，“green goat”)，命名的反应时要短于无关条件(例如，“green rug”) (Damian & Dumay, 2007, 2009)。采用图画-词汇干扰范式(Damian & Martin, 1999; Meyer & Schriefers, 1991)和掩蔽启动范式(Forster & Davis, 1991; Kinoshita & Woollams, 2002; Malouf & Kinoshita, 2007; Schiller, 2008)研究者在印欧语系中均发现了类似的结果。

最近，研究者对汉语口语词汇产生的单词形式编码过程中音节和音素的作用存在不同看法。汉语作为一种声调语言与字母语言存在很大的不同。首先，相比英语、荷兰语，汉语的

---

[2] 注：音素(phoneme)是一个语音系统中能够区别意义的最小语音单位。按照发音特质可以分为两类，元音(Vowel，简称 V)和辅音 (Consonant，简称 C)。

[3] 注：音节(syllable)在语音学中被用来指称两个最小响度之间有一个最大响度的一连串语音。在语音分析中，音节被定义为每种语言都具有的一个单位，是辅音和元音组成的序列。所有语言都有 CV 和 CVV 音节结构例如在汉语中/la1/是一个 CV 结构的音节，其中/l/是辅音，/a/是元音，而/lai2/是一个 CVV 结构的音节，包括一个辅音/l/和两个元音/a/与/i/。

音节数量很少，不计声调有 400 个音节左右，计声调也只有 1200 个左右，而荷兰语中大约有 12000 个音节(张清芳，杨玉芳，2005; Levelt et al., 1999)。其次，汉语中的音节边界十分清晰，不存在重新音节化现象，而在拼音语言的口语产生中，会出现大量跨越音节边界重新组合成音节的现象，根据 Levelt 等(1999)的观点，重新音节化现象是心理词典中不存储音节的主要原因，而在汉语中更为经济的方式是将所有音节存储起来，并在音韵编码过程中直接提取。因此汉语的语音特点决定了汉语的主要音韵编码单元更可能是音节，而非音素。

J.-Y. Chen(2000)分析汉语语误语料库发现音节交换错误的发生率显著高于音素交换错误，表明音节在汉语口语产生中的重要性高于音素。J.-Y. Chen等(2002)利用内隐启动范式探索汉语双音节词汇产生中的音韵编码阶段，发现了音节相同条件下的启动效应，但音素相同条件下无任何效应，这表明无声调的音节在音韵水平上能作为一个独立的计划单元。研究者采用掩蔽启动范式(J.-Y. Chen et al., 2003; J.-Y. Chen et al., 2016)和图画-词汇干扰范式(张清芳，杨玉芳，2005)得到了同样的结果。You, Zhang和Verdonschot(2012)利用图画命名和单词命名两种任务均发现了音节启动效应，为汉语口语中音节是音韵编码单元提供了更强的证据。岳源和张清芳(2015)采用图画-词汇干扰范式，通过比较即时命名、延迟命名以及延迟命名与发音抑制任务的结合,考察了汉语口语产生中音节和音段(segments)[4](多个音素结合形成的一个单元，其音素个数少于音节)在单词形式编码的不同阶段所产生的效应，结果表明音节的促进效应发生在音韵编码阶段，证明了音节是音韵编码过程的合适单元，音段可能在随后的语音编码阶段起作用。

与此不同，有一些研究发现了粤语口语产生中音段的促进效应。Wong和H.-C. Chen(2008, 2009)采用图画-词汇干扰范式，在粤语中发现当干扰词与目标词的韵母和声调相同时，图画命名显著快于无关条件，因此他们认为粤语口语产生中的主要音韵编码单元可能是音段(同见Wong, Huang, & H.-C. Chen, 2012)。与行为结果不一致的是，Wong，Wang, Ng和Chen(2016)利用图画-词汇干扰范式与ERP技术发现粤语口语产生中的音节效应出现在500~650 ms之间。Wong等(2012)认为粤语的口语产生过程可能与汉语普通话存在差异，这是因为粤语的音韵体系比汉语普通话复杂：例如，普通话中仅有两种尾音/n/和/ng/，而粤语中包含六种尾音/p/, /t/, /k/, /m/, /n/和/ng/，而且/ng/在粤语中还可以作为首音出现。粤语中包括了6个声调，而普通话中仅包括4种声调。

[4]注：音段(segment)在汉语中指的是大于音素而小于音节的组合单元(Wong & H. -C. Chen, 2008, 2009)， 例如，对于一个 CVC 结构的汉语音节 /lan2/， 可以划分为两个音段， 首 CV 结构的/la/和尾音段 VC 结构的/an/。

研究者也采用高时间分辨率的ERP技术考察了音节和音素效应发生的时间进程。Qu等(2012)采用首音素重复范式(如首音素重复的“黄盒子”和首音素不重复的“绿盒子”)，在图画呈现后的200~300 ms之间发现了首音素重复效应，表明音素信息在音韵编码阶段也会被激活(同见Yu, C. Mo, & M. Mo, 2014)。O’Seaghdha等（2013）指出在Qu等(2012)发现的ERP波形差异可能仅仅反映了颜色词和目标词在首音素相同时的联结，并不能表明音素是音韵编码的单元。另一方面，Qu等(2012)的研究在英国完成，被试为留学英国的学生，二语水平较高。研究发现第一语言的加工过程会受到第二语言学习迁移的影响(Jones et al., 2012; Verdonschot, Nakayama, Zhang, Tamaoka, & Schiller, 2013)。Verdonschot等(2013)选取熟练英-汉双语者完成英语和汉语两种语言的启动任务，发现了汉语任务中的启动词与目标词音节结构相同时的首音素启动效应，表明第二语言会影响第一语言的加工过程。最近，研究者采用图-图干扰范式与延迟图画命名任务，在目标图画呈现后的200~400 ms与400~600 ms发现了音节启动效应，却未发现任何音段启动效应(Wang et al., 2017)。Zhang和Damian(2019)采用掩蔽启动范式，首次在汉语普通话的口语词汇产生过程中同时探测到了音节效应(300~400 ms)和音段效应(500~600 ms)，基于元分析中各个阶段对应的时间进程(Indefrey, 2011)，研究者认为这两个效应可能分别发生在音韵编码阶段和语音编码阶段。

Roelofs(1997)提出的WEAVER(Word-form Encoding by Activation and VERification)模型是基于印欧语系的结果建立的，该模型假设音韵编码的单元是音素，单个音节内部的音韵编码是一个增长式的编码过程，从音节的首音开始到核心元音最后到尾音(Levelt et al., 1999)。根据汉语和印欧语系关于音韵编码单元的不同研究结果，O’Seaghdha等 (2010)提出了合适单元假说(Proximate units Principle)来解释跨语言间音韵编码单元的不同(如图1所示)。合适编码单元(Proximate units)指的是激活词汇的词素之后首先被激活的音韵编码单元，这一假设认为音韵编码单元中最先选择的单元存在语言上的差异，印欧语系如英语或荷兰语中最先选择的单元是音素，而在汉语中则为音节。在印欧语系中，讲话者在选择音素后，结合节律信息进行音节化过程，从心理音节表中提取音节准备发音运动程序。汉语口语产生中讲话者则在选择音节后进一步分解为音素或音段信息(音韵编码阶段)，准备发音运动程序(语音编码过程)，最后进行发音输出口语产生的结果(发音阶段)。Roelofs (2014)利用计算机模拟的方法将合适单元假说加入到了WEAVER模型的参数设定中，验证了合适单元假说。

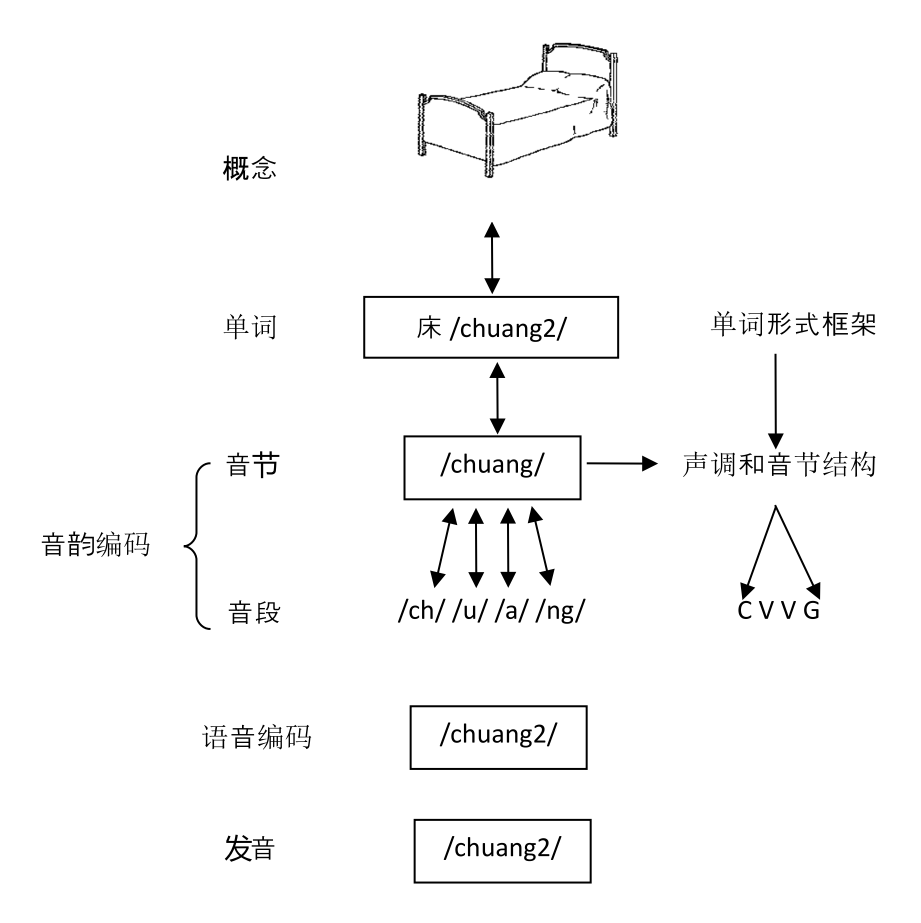


图 1 O′Seaghdha 等(2010)提出的汉语单音节词汇产生的示意图

尽管如此，我们可以看到有关音节和音素在汉语口语产生中的作用及其发生阶段仍存在激烈争论。已有研究存在行为与 ERP 结果的不一致(Wong & H.-C. Chen, 2008,2009; Wong et al., 2016)，以及 ERP 结果之间的不一致(Qu et al., 2012; Wang et al., 2017; Zhang & Damian, 2019)。在两个发现音素效应的研究中(Qu et al., 2012; Yu et al., 2014)，均未将音素与音节效应进行比较，且存在被试的加工方式可能受到第二语言迁移的影响。已有 ERP 研究均采用了图画命名任务，包括了概念准备、词条选择、音韵编码、语音编码和发音等一系列加工过程，这些过程之间会产生相互影响，比如概念准备对词汇选择和音韵编码的影响。在 ERP 研究中，音节条件下的激活较强，而音素条件下的激活较弱，难以探测到(Wang et al., 2017)，这提示研究者应该采用对音素效应更为敏感的实验任务。

在口语产生领域，研究者采用内隐启动范式对音韵编码进行了深入研究。在这一任务中，被试先学习几个词对，比如“海豹-夕阳，记忆-媳妇，皇帝-喜事，贝壳-细胞”。这四个单词对中的第二个单词第一个字的音节是相同的，称之为同质(homogeneous)条件。对照组单词对的第二个单词在音韵上毫无关联，称之为异质(heterogeneous)条件。在测试阶段呈现单

词对中的第一个词(线索词如“海豹”)，被试要说出对应的第二个词(目标词如“夕阳”)(张清芳，2008)。在内隐启动范式中，人们在目标词之间有音韵联系的情况下会产生准备效应，出现音韵促进效应。在印欧语系中，研究者采用内隐启动范式灵敏地探测到了单个音素的准备效应(Meyer, 1990, 1991; Damian & Bowers, 2003)，表明该范式对音韵编码单元是敏感的。在内隐启动范式的每组测试中，被试仅需要针对四个线索词说出目标词，且在测试之前被试已经多次学习了这四个词对，能在线索词呈现时快速说出目标词。研究者认为每组测试中的目标词语在测试过程中都是从工作记忆中提取出来，不涉及到概念的加工(Meyer, 1990, 1991; J.-Y. Chen et al., 2002; Cholin, Schiller, & Level, 2004)，提取后进入音韵编码过程。相对于图片命名任务来说，内隐启动任务中不涉及到口语词汇产生中的概念准备阶段。根据此分析，我们猜测内隐启动范式可能对音素效应更为敏感。

本研究将内隐启动范式与脑电技术相结合，考察音素和音节效应的时间进程，而且我们选取了低二语水平且日常生活中均不接触英语的被试，以排除第二语言对第一语言的影响。根据合适单元假说，我们预测音节效应发生的时间早于音素效应，音节效应发生在音韵编码阶段，而音素效应可能出现在音韵编码阶段后期或语音编码阶段。

2 方法

2.1 被试

20 名成年人(年龄 22 岁至 46 岁，16 名男性)，学历为初中 2 人，高中 3 人，大专 6 人，本科 9 人。所有被试日常生活中均不接触英语，且都是在初中阶段才开始学习英语。在实验开始前所有被试都填写了英语水平自我评定调查问卷，按照 0~7(0 代表完全不会，7 代表非常好)标准自评，所有被试的评分均值为阅读 1.53(SD = 0.94，范围 0~4)分，书写 1.37(SD = 1.18，范围 0~3)分，口语 1.42(SD = 1.35，范围 0~3)分，听力 1.32(SD = 0.98，范围 0~3)分。所有被试无脑部疾病史，母语为汉语，普通话标准，视力或矫正视力正常。在参加实验之前被试阅读知情同意书并签字，实验之后获得一定报酬。

2.2 材料

实验材料包括 36 对仅有语义联系(无任何正字法和音韵相关)的双字词对(例如：神仙-凡人，“神仙”为线索词，“凡人”为目标词)，其中 32 对为正式实验材料，4 对为练习材料。32 对词汇根据目标词之间的相关关系分为 4 套音节相同(但声调不同)和 4 套首音素相同的词汇，每套包括 4 对词汇，分别构成 4 套音节同质条件和 4 套音素同质条件。从音节同质条件下的 4 套词汇中均选择 1 个词对，这样可以形成 4 个目标词无任何语音联系的词对，

共组成 4 套音节异质条件。用同样的方法形成 4 套音素异质条件(实验材料见附录 1)。这样所形成的音节同质条件和异质条件、音素同质和异质条件下的词对是相同的，唯一的不同是目标词的音韵线索不同。

为了匹配音节和音素条件中所用词对中线索词和目标词之间的语义关联度，我们招募了 21 名中国人民大学的学生，要求他们在 5 点量表上进行语义关联度评定，1 代表完全无关，5 代表完全相关，结果发现音节材料(M = 4.02, SD = 0.54)与音素材料(M = 4.12, SD = 0.52)的语义关联度没有显著差异，$F(1, 20) = 1.97$, $p = 0.176$。

为了匹配音节和音素条件中同质组与异质组内目标词之间的语义关联度，我们招募了 20 名中国人民大学的学生，要求他们在 5 点量表上对两个目标词之间(例如“危机-围墙”、“危机-卫星”或“围墙-尾巴”等)的语义关联度进行评定，1 代表完全无关，5 代表完全相关，结果发现在音节材料中，两组目标词之间的语义关联度都很低，且同质组(M = 1.34, SD = 0.29)与异质组(M = 1.32, SD = 0.35)的语义关联度没有显著差异，$F(1, 19) = 0.177$, p = 0.862。在音素材料中，同质组(M = 1.28, SD = 0.29)与异质组(M = 1.27, SD = 0.32)的语义关联度没有显著差异，$F(1, 19) = 0.18$, $p = 0.79$。同时我们也对同质组与异质组中线索词之间的语义关联度进行了测量，结果显示，在音节材料中，两组线索词之间的语义关联度都很低，且同质组(M = 1.41, SD = 0.36)与异质组(M = 1.33, SD = 0.33)的语义关联度没有显著差异，$F(1, 19) = 1.71$, $p = 0.103$。在音素材料中，同质组(M = 1.24, SD = 0.24)与异质组(M = 1.31, SD = 0.27)的语义关联度没有显著差异，$F(1, 19) = 1.32$, $p = 0.23$。

2.3 设计

实验设计包括了目标词相关条件(同质和异质条件)和相关类型(音节和音素)的两因素完全重复测量设计，两个自变量均为被试内因素。每个被试重复相同的程序四次，每次都包括一个同质条件组和一个异质条件组，同异质条件出现的顺序在被试间进行平衡。在同质和异质条件组中，每套中的四对单词重复出现四次，共包括 16 次测试。对于每次重复，每组条件，每套线索词出现的顺序都是随机的，并保证相同的单词不会连续呈现。每个被试要完成 256(2 种相关条件×2 两种相关类型×4 套材料×4 对单词×4 次重复)次测试。

2.4 仪器

实验程序由 E-Prime 2.0 编制，PST SRBOX 反应盒，麦克风和计算机。图片都呈现在 PIII-667 计算机屏幕中央。被试的反应通过 PSTSR-BOX 连接的麦克风记录。实验材料的呈现、计时和被试的反应时由电脑控制与收集。主试记录被试的反应正确与否。

2.5 程序

实验分为三个阶段，学习阶段、测试阶段和正式实验阶段。在学习阶段，被试学习音节相同或音素相同条件下的16组词对加上练习用的4组词对，共20组词对，开始时屏幕左右两边分别呈现提示词和目标词，被试可以自己控制学习的时间，要求熟记全部20组词对。学完之后进入测试阶段，在屏幕中央呈现刚才被试学过的词对中的第一个词(线索词)，字体大小为30号宋体字，被试要大声清晰地说出对应于该词的第二个词(目标词)。被试反应后屏幕上呈现正确答案。在测试阶段每对词对重复出现三次。如果被试出现错误，测试阶段结束后，对所有错误的词对再进行一次学习，然后测试，直至被试准确熟练地记住全部词对。

在实验阶段，首先呈现注视点“+”500ms，之后呈现300ms空白屏，接着屏幕中央呈现线索词(不超过3000ms)，呈现字体大小为30号宋体字。被试的任务是看到提示词之后尽量准确快速地说出所对应的目标词，主试在被试说出目标词后按键判断反应正误，一次测试结束后间隔1500 ms后开始下一次测试。

2.6 EEG 记录和分析

脑电信号经由Neuroscan公司生产的ESI-64导脑电记录系统进行采集，64导银、氯化银电极以国际通用的10-20方式固定于电极帽上。左侧乳突作为参考电极，额头中央连接接地电极。位于左眼上下眶的电极记录垂直眼电(VEOG)，位于左右眼角外1cm处的电极记录水平眼电(HEOG)。电极与头皮之间的阻抗小于5kΩ。信号经放大器放大，连续记录时滤波带通为0.05~100Hz，采样频率为500 Hz。

对记录的EEG数据进行离线处理，选择双侧乳突作为重新参考电极(Wang, Bastiaansen, Yang, & Hagoort, 2011)。通EEGLAB软件进行眼电伪迹的校正。数据分析时滤波带通为0.1~30 Hz，按照目标图呈现前200ms和呈现后600ms的时间段对脑电进行分段，前200ms作为基线进行基线校正，删除眼电、肌电等伪迹。在伪迹校正中将波幅超过±100 μV去除。

根据矢状轴和冠状轴选取6个兴趣区(Regions of Interest, ROI)，其中每个外侧ROI的电压值是三个电极点的平均波幅，例如，左前(F5, F7, FC5)，中前(Fz, FCz, Cz)，右前(F6, F8, FC6)，左后(P5, P7, CP5)，中后(CPz, Pz, POz)，右后(P6, P8, CP6)。以各时段的平均波幅为因变量，进行2(相关条件：同质、异质) × 2(相关类型：音节、音素) × 2(脑区：前、后) × 3(半球：左、中、右)重复测量方差分析。当球形假设不成立时，方差分析的p值均用 Greenhouse-Geisser校正。

3 结果

20 名被试中有 2 名被试由于伪迹太多被删除，共分析了 18 名被试数据，每名被试每个条件下的有效叠加试次均超过 40 次。

3.1 行为结果

删除命名错误的数据(0.52%)，其它声音比如 “嗯”或“啊”引起的反应及话筒未反应(1.19%)以及偏离平均值 2.5 个标准差以外的数据(5.41%)。方差分析结果表明在音节类型中，同质组(M = 844ms, SD = 87)与异质组(M = 850ms, SD = 90)的反应时无显著差异，$F_1(1, 17) = 0.737$，$p = 0.406$, $F_2(1, 15) = 0.113$, $p = 0.741$；在音素类型中，同质组(M = 791ms, SD = 89)与异质组(M =788ms, SD = 101)的反应时无显著差异，$F_1(1, 17) = 0.229$, $p = 0.638$，$F_2(1, 15) = 0.219$，$p = 0.647$。

3.2 ERP 结果

反应时小于 600 ms 的数据(5.2%)和大于 1500ms 的数据(2.7%)被剔除，仅有反应正确的试次纳入下面分析。

根据已有研究(Qu et al., 2012; Zhang & Damian, 2019)，我们确定了 6 个连续的时间窗进行平均波幅分析，分别为 0~100 ms、100~200 ms、200~300 ms、300~400 ms、400 ~500 ms 、500~600 ms。对每个时间窗口的平均波幅进行 2(相关类型：音节、音素)×2(相关条件：同质、异质)×2(脑区：前、后) × 3(半球：左、中、右)重复测量方差分析，当球形假设不成立时，方差分析的 p 值均用 Greenhouse-Geisser 校正。同时用 R 软件中的 fdrtool 对 p 值进行了 *fdr* 校正(见表 1)。

表 1 相关类型、条件、脑区和半球为自变量在 0~600 ms 中方差分析的结果

变异	时间窗口（ms）					
	0~100	100~200	200~300	300~400	400~500	500~600
条件(1，17)	—	—	—	4.21†	—	—
类型(1，17)	—	—	—	5.09*	—	—
条件×类型(1，17)	—	—	—	6.31*	—	—
条件×脑区(1，17)	—	7.00*	5.34*	—	—	4.49†
条件×半球(2，34)	—	—	—	—	—	—
条件×类型×脑区(1，17)	—	4.01†	—	—	—	—
条件×类型×半球(2，34)	—	—	—	—	—	—
条件×类型×半球×脑区(2，34)	—	—	—	4.30*	—	—

注：$^{\dagger}p < 0.1$, * $p < 0.05$, ** $p < 0.01$, p 值均经过 FDR 校正。

在 100~200 ms 时间窗口内, 相关类型、相关条件和脑区的交互作用边缘显著，$F(1, 17)$ =

4.01, p = .061, η_p^2 = 0.191，如表 2，进一步配对分析表明，音节同质条件与异质条件相比有显著差异，表现为同质条件的平均波幅更正(见图 2)。两个兴趣区上的差异相比达到了显著水平(经 *fdr* 校正)(见图 3)，左前区：t(17) = 3.02, p = .024；右前区：t (17) = 3.35, p = .02; 6 个兴趣区内均未发现任何音素效应, *ps* > .462。

在 200~300 ms 时间窗口内，未发现相关类型和相关条件的任何交互作用, *ps* > .105。但进一步兴趣区配对分析表明，如表 2，音节同质条件与异质条件相比有显著差异，表现为同质条件的平均波幅更正(见图 2)。在三个兴趣区上的差异相比达到了显著水平(经 *fdr* 校正)(见图 3)，左前区：t(17) = 2.55, p = .032；中前区：t(17) = 2.19, p = .043；右前区：t(17) = 3.82, p = .012；6 个兴趣区内均未发现任何音素效应, *ps* > .230。

表 2 不同时间窗口的音节和音素效应

兴趣区	时间窗口(ms)					
(1,17)	0~100	100~200	200~300	300~400	400~500	500~600
			音节效应			
左前	—	3.02*	2.55*	2.38*	—	—
中前	—	2.10†	2.19*	2.63*	—	—
右前	—	3.35*	3.82*	3.26*	—	—
左后	—	—	—	2.04†	—	—
中后	—	—	—	2.28*	—	—
右后	—	—	—	2.67*	—	—
			音素效应			
左前	—	—	—	—	—	—
中前	—	—	—	—	—	—
右前	—	—	—	—	—	2.71*
左后	—	—	—	—	—	—
中后	—	—	—	—	—	—
右后	—	—	—	—	—	2.43*

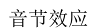
注：$^{\dagger}p < 0.1$, * $p < 0.05$, ** $p < 0.01$, p 值均经过 FDR 校正。

在 300~400 ms 的时间窗口内，相关类型和相关条件的交互作用显著，F(1, 17) = 6.31，p = .022，η_p^2 = 0.271，如表 2，进一步配对分析表明，音节同质条件与异质条件相比有显著差异，表现为同质条件的平均波幅更正(见图 2)。在五个兴趣区上的差异相比达到了显著水平(经 *fdr* 校正)(见图 3)，左前区：t(17) = 2.38, p = .035；中前区：t (17) = 2.63, p = .031；右前区：t (17) = 3.26, p = .02；中后区：t (17) = 2.28, p =.039；右后区：t(17) = 2.67, p = .031。6 个兴趣区内均未发现任何音素效应, *ps* > .315。

在 400~500ms 的时间窗口内, 未发现相关类型和相关条件的任何交互作用, *ps* > .271。

在 500~600ms 的时间窗口内，未发现相关类型和相关条件的任何交互作用, *ps* > .233。

但进一步兴趣区配对分析表明，音素同质条件与异质条件相比有显著差异，表现为同质条件的平均波幅更正(见图 2)。在两个兴趣区上的差异相比都达到了显著水平(经 *fdr* 校正)(见图 3)。右前区：$t\,(17) = 2.71, p = .031$；右后区：$t\,(17) = 2.43, p = .035$。6 个兴趣区内均未发现任何音节效应, $ps > .227$。



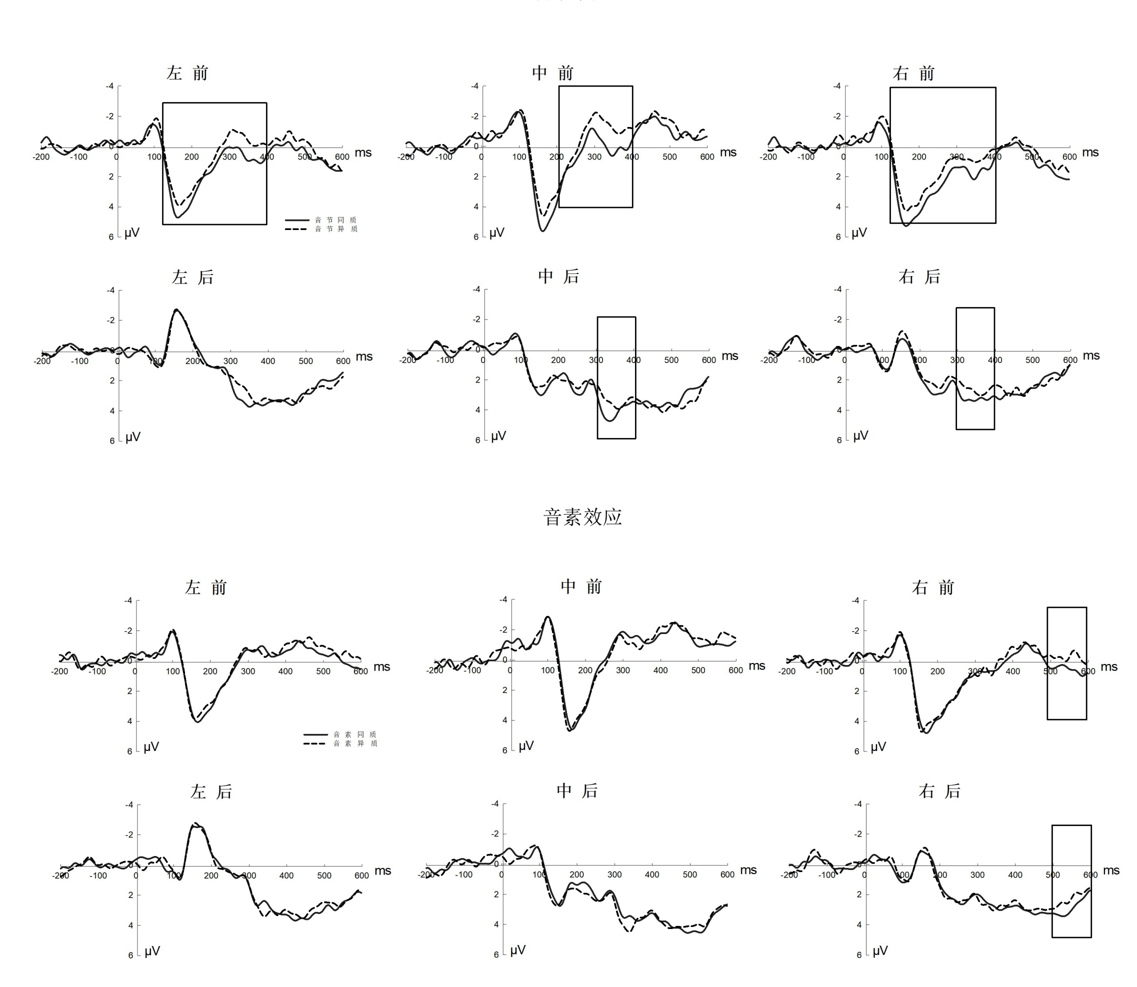


图 2 6 个兴趣区内音节和音素效应的平均波形图(方形框表示两类条件对比有显著差异)

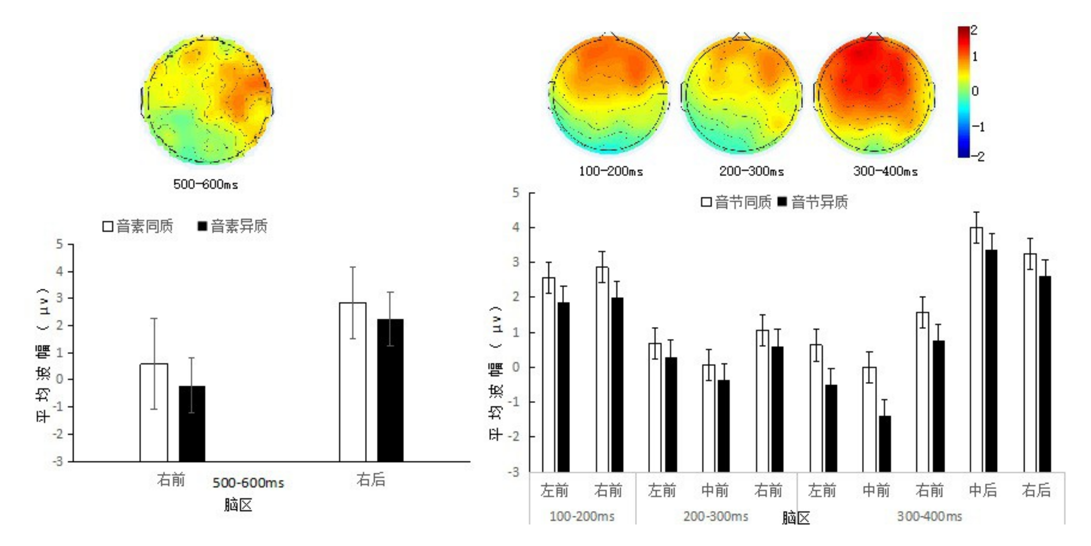


图 3 在不同时间窗口下的音节和音素效应的平均波幅与各个时间窗口下的地形图(注：图中误差棒均为 95%CI)

4 讨论

本研究采用内隐启动范式，考察了汉语口语产生过程中音韵编码阶段的加工单元，并比较了音节和音素提取的时间进程。行为结果未发现音节效应和音素效应。ERP 平均波幅的分析发现，音节效应出现在线索词呈现后的 100~400 ms 之间，最先出现差异的区域是左右的前部电极，随后中前区出现差异，最后在脑前部区域以及右中后部区域都出现了显著差异。与音节效应相比，音素效应仅出现在线索词呈现后的 500~600 ms 之间，分布于右侧前部和后部。波幅的差异都表现为相关条件比无关条件诱发的波形更正。这些发现表明在内隐启动任务中，词汇的音节信息首先被提取，之后是音素信息，与合适编码单元假说(O'Seaghdha et al., 2010)的观点一致，在词汇口语产生过程中，词汇被提取后，人们首先会在音韵编码阶段提取词汇相应的音节，而不是音素。

在行为结果上，本研究未发现汉语口语产生中的音节效应，与这与已有采用相同任务的研究结果不一致(J.-Y. Chen et al., 2002; 张清芳, 2008)。一个可能的原因是实验材料方面的问题，为了确认实验材料的选取是否合适，我们选择了 6 名未参加正式实验的中国人民大学的研究生完成实验，结果发现 6 名学生都出现音节启动效应，音节同质组(M = 745ms, SD = 153)比音节异质组(M = 765ms, SD = 157)的反应更快，$t(5) = 3.10$, $p = .022$，而未发现音素启动效应，$p > 0.05$。这与已有研究结果一致，表明实验材料的选择是有效的。值得注意的是，已有的研究均选择了大学生或研究生作为被试，这些人日常生活中都会大量地接触英语材料，这在一定程度上表明受教育程度和二语水平可能会影响母语口语产生过程。事实上，本研究中的被试也表现出了音节同质组的反应时短于音节异质组，出现了微弱的促进效应。相比而言，研究生组对相同的实验材料和实验任务则出现了 20 ms 左右的促进效应，表明

研究生组对于音节信息比本研究中的被试组更为敏感。此外，研究发现实验组产生目标词的平均潜伏期为 840 ms 左右，而材料测试组的平均潜伏期为 750 ms 左右，在内隐启动范式中，被试需要事先记忆线索词和目标词的联系，可能由于被试对词对的联系记忆不如材料测试组，当反应时延长时被试在同质组和异质组产生的差异变得不显著。这一对比提示被试的受教育程度以及日常与文字等材料接触的经验可能会影响讲话者的口语产生过程。关于二语水平和受教育程度对口语产生过程的影响机制需要进一步考察。

在内隐启动范式中，研究者一般不会去匹配线索词和目标词的词频和笔画数等，研究者主要关注的是不同音韵相关条件下，同质与异质相比是否存在显著差异。因为所用实验材料的不同，可能会出现音节条件和音素条件下反应时存在差异，但这并不影响对音节效应(音节同质–异质)和音素效应(音素同质–异质)的考察。大量的行为实验已经验证了这一实验逻辑的可靠性(J.-Y. Chen et al., 2002; Cholin et al., 2004; Damian & Bowers, 2003; Meyer, 1990, 1991; 张清芳，2008)。例如，J.-Y. Chen 等(2002)采用内隐启动范式，对汉语词汇产生过程进行研究，发现音节同质条件下的平均反应时在不同的音节和不同被试下分别为 658 ms(实验 2)，666 ms(实验 3)，662 ms(实验 4)，而音素同质条件下的反应时为 622 ms (实验 5)。张清芳(2008)采用同样的实验任务发现音素同质平均反应时为 620 ms，而音节同质的平均反应时为 633 ms。可以看到，平均反应时之间是存在较大差异的，且均表现为音节同质条件长于音素同质条件，与本实验的发现一致，表明实验材料的选择确实会造成反应时上的差异，但并不影响音节或音素效应的获得，研究者在做比较时并不关注不同材料之间的差异。

本研究的 ERP 结果显示在内隐启动范式中，讲话者对音节的提取先于音素。这一结果与已有采用掩蔽启动范式的 ERP 结果一致(Zhang & Damian, 2019)，音节效应发生在 300~400 ms，根据对图画命名过程的元分析结果(Indefrey, 2011)，研究者认为这音节效应发生在音韵编码阶段。Wang 等(2017)采用图-图干扰范式和延迟图画命名任务(粤语)，在目标图画呈现后的 200~400 ms 与 400~600 ms 发现了音节启动效应，但未发现音素效应，研究者将这两个时间窗口的效应分别解释为发生在音韵编码和语音编码阶段。已有口语词汇产生的时间进程都是基于图画命名任务，图画命名包括了概念准备、词汇选择、音韵编码、语音编码和发音过程，内隐启动任务中的产生过程不包括概念准备，是从词汇选择开始到发音过程。因此，与图画命名范式相比，内隐启动范式中的音韵编码阶段会相对较早，这与我们发现的音节效应在 100 ms 左右就开始的结果是一致的，表明被试在选择完词汇后首先提取了音节，且音节效应一致持续到 400 ms，分布的脑区从脑区前部扩散至右侧中后部电极。因

此，我们认为本研究发现的音节效应发生在音韵编码阶段。

音素效应出现在线索词呈现后的 500~600 ms 之间，发生在音节效应之后。根据 O’Seaghdha 等(2010)提出的模型(见图 1)，在汉语口语产生过程中，音韵编码阶段音节被提取后会进一步分解为各个音素，在此之后各个音素的提取是平行进行的(Roelofs, 2015) 或者是以从左至右的顺序提取的(J.-Y. Chen et al., 2016)。同时人们也会提取相应的节律信息，形成单词形式的框架，将音素插入到相应的框架中去，进行语音编码为发音做好准备。音素效应的可能发生在音韵编码或者语音编码阶段。第一种可能性是音素效应发生在音韵编码阶段，在音节提取之后将其分割为各个音素，这一过程发生在音韵编码阶段的后期。第二种可能性是音素效应发生在语音编码阶段，Zhang 和 Damian(2019)在图画命名过程中发现了 500~600 ms 时间窗口内的音段效应，与图画命名元分析的时间进程对比，他们认为这一效应可能出现在语音编码阶段，反映了音段的启动作用。与已有的研究相比，Wang 等(2017)的研究同时设置了音节和音素相关条件，仅仅发现了音节效应，未探测到音素效应；Qu 等(2012)和 Yu 等(2014)的研究中仅仅设置了音素相关条件，不能比较音节和音素效应发生的相对时间进程。本研究无法区分这两种解释，需要通过研究任务之间的对比(即时命名、延迟命名，以及延迟命名和发音抑制的结合)对这两种解释进行考察。

本研究发现音素效应的时间窗口晚于 Qu 等(2012)发现的 200~300 ms 和 Yu 等(2014)发现的 180~300 ms，他们将其解释为音韵编码阶段音素的激活。在图画命名的时间进程中，当图画命名潜伏期为 600 ms 左右时，180~300 ms 的时间窗口对应于词汇选择阶段(Indefrey，2011)，音韵编码应该发生在 275~400 ms 的时间窗口内。Qu 等(2012)的研究中命名反应时为 976 ms，但音素效应却出现得很早，O'Seaghdha 等(2013)指出 Qu 等(2012)的发现可能反映了形容词和目标词在音素相同时的音韵联结，而不是反映了在音韵编码阶段音素作为提取单元的激活，这一质疑仍然需要进一步的实验验证。也有可能是所采取的实验任务不同导致了音素效应的时间窗口不同。此外，已有研究在脑区之间比较音素相关和无关条件时并未进行统计校正(Qu et al., 2012; Yu et al., 2014; Wong et al., 2016; Wang et al., 2017)，我们尝试采用 *fdr* 校正对已有研究报告的 p 值进行校正后发现，基本上所有的 p 值在校正后都不显著($p > 0.05$)。本研究中的实验数据在统计校正后仍然显著，能在一定程度上避免这一问题，表明在内隐启动范式下发现了显著的音素效应。

本研究发现无论是音节效应还是音素效应，在波形上都表现为同质条件比异质条件更正，这与其他 ERP 的研究发现一致(Qu et al., 2012; Yu et al., 2014; Wong et al., 2016; Wang et

al., 2017; Verdonschot, Tokimoto, & Miyaoka, 2019)，反映了音节或音素相同时促进了口语产生过程。同时本研究中所发现的音节促进效应位于线索词出现后的 100~400 ms 之间，这与经典的 P2 和 P3 波形所出现的时间窗口相似(Hackley, Woldorff, & Hillyard, 1990; Donchin & Coles, 1988)。研究表明 P2 成分与视觉特征的探测有关，而且 P2 成分对正字法和音韵加工任务、语义分类任务和词汇分类任务中不同条件的设置比较敏感，会出现波幅上的差异 (Luck & Hillyard, 1994)。研究也发现 P3 的波幅大小与注意的程度有关(Gray & Burgess, 2004)，也可能与决策做出后的认知闭合(cognitive closure)(Desmedt, 1980; Verleger, 1998)或与有意识的信息通达过程相关(Picton, 1992)。这一成分包括了两个子成分(P3a 和 P3b)，P3a 分布在前额叶部位，当有未预期的事件出现时引发这一成分，反映了自动的注意调节机制；P3b 分布在顶枕区域，与注意、工作记忆和复杂的认知加工(包括语言加工)相关。这提示与音节异质条件相比，音节同质条件下可能诱发了被试更多的注意，促进了被试的口语产生过程。关于注意如何调节口语产生中的音韵编码过程以及波形的变化，则需要更进一步的研究。

目前，口语词汇产生领域的 ERP 研究一般都采用实验条件与基线条件对比的设计，比较在各个时间窗口下两类条件是否达到了显著差异，时间窗口的选择具有较大差异，且由于实验任务和被试不同，在比较不同研究之间的结果时要尤为谨慎。尽管如此，不同研究结果所得到两种效应的时间序列模式是可以对比的，我们的研究发现音节效应的发生早于音素效应，与采用掩蔽启动范式所得到的模式一致(Zhang & Damian, 2019)。

5 结论

综上，我们首次采用内隐启动范式，利用词语作为目标材料，选择英语水平低的汉语母语者，排除作为二语的英语的音韵编码单元可能对汉语口语词汇产生过程的影响后，仍然发现了音节效应(100~400 ms)早于音素效应(500~600 ms)的模式，表明在汉语口语词汇产生过程的词汇选择之后，讲话者会首先提取音节信息，然后将音节分解为音素，支持了合适编码单元假说。

参考文献

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*, 464–475.

Chen, J.-Y. (2000). Syllable errors from naturalistic slips of the tongue in Mandarin Chinese. *Psychologia: An International Journal of Psychology in the Orient*, *43*, 15–26.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language*, *46*, 751–781.

Chen, J.-Y., Lin, W.-C., & Ferrand, L. (2003). Masked priming of the syllable in Mandarin Chinese speech production. *Chinese Journal of Psychology*, *45*, 107–120.

Chen, J.-Y., O'Séaghdha, P. G., & Chen, T.-M. (2016). The primacy of abstract syllables in Chinese word production. *Journal of Experimental Psychology:Learning, Memory, and Cognition*, *42*, 825–836.

Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language, 50*(1), 47-61.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*, 119–132.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*, 195–209.

Damian, M. F., & Dumay, N. (2009). Exploring phonological encoding through repeated segments. *Language and Cognitive Processes*, *24*, 685–712.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 345–361.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production.*Psychological Review*, *93*, 283–321.

Desmedt, J. E. (1980). P300 in serial tasks: An essential post-decision closure mechanism. *Progessing in Brain Research*, *54*, 682-686.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of cognitive updating? *The Behavioral and Brain Sciences*, *11*, 357-427.

Forster, K.I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, *30*, 1–25.

Gray, J. & Burgess, G. C. (2004). Personality differences in cognitive control? BAS, processing efficiency, and the prefrontal cortex. *Journal of Research in Personality*, *38*, 35-36.

Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1990). Cross-modal selective attention effects on retinal, myogenic, brainstem and cerebral evoked potentials. *Psychophysiology*, *27*, 195-208.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 1–16.

Jacobs, C. L., & Dell, G. S. (2014). “hotdog”, not “hot” “dog”: The phonological planning of compound words. *Language, cognition and neuroscience*, *29*, 512–523.

Kinoshita, S., & Woollams, A. (2002). The masked onset priming effect in naming: Computation of phonology or speech planning. *Memory and Cognition*, *30*, 237–245.

Levelt, W. J. M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291-308.

Malouf, T., & Kinoshita, S. (2007). Masked onset priming effect for high-frequency words: Further support for the speech-planning account. *Quarterly Journal of Experimental Psychology*, *60*, 1155–1167.

Meyer, A. S. . (1990). The time course of phonological encoding in language production: the encoding of successive syllables of a word ☆. *Journal of Memory & Language, 29*(5), 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *17*, 1146–1160.

O’Seaghdha, P. G.,Chen, J.-Y., & Chen, T.-M. (2010). Proximate units in word production: phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition*, *115*, 282–302.

O'Seaghdha, P. G., Chen, J.-Y., & Chen, T.-M. (2013). Close but not proximate: The significant of phonological segments in speaking depends on their functional engagement. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, E3.

Parker Jones, O., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., . . . Price, C.J. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb Cortex, 22*(4), 892-902.

Picton, T. W. (1992). The P300 wave of the human event related potential. *Clinical Neurophysiology Lisse*, *9*(1), 456-479.

Qu, Q., Damian, M. F. & Kazanina, N. (2012). Sound-sized segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 14265–14270.

Qu, Q., Zhang, Q., & Damian, M. F. (2016). Tracking the time course of lexical access in orthographic production: An event-related potential study of word frequency effects in written picture naming. *Brain and Language*, *159*, 118–126.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.

Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*, *57*, 22–37.

Schiller, N. O. (2008). The masked onset priming effect in picture naming. *Cognition*, *106*, 952–962.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Erlbaum.

Verdonschot, R. G., Nakayama, M., Zhang Q., Tamaoka K., & Schiller, N. O. (2013). The proximate phonological unit of Chinese-English bilinguals: Proficiency matters. *PloS One*, *8*(4):e61454.

Verleger, R. (1988). Towards an integration of P3 research with cognitive neuroscience. *Behavioral and Brain Sciences*, *21*, 150-154.

Verdonschot, R. G., Tokimoto, S., & Miyaoka, Y. (2019). The fundamental phonological unit of Japanese word production: An EEG study using the picture-word interference paradigm. *Journal of Neurolinguistics*, 184-193.

Wang, L., Bastiaansen, M., Yang, Y. F., & Hagoort, P. (2011). The influence of information structure on the depth of semantic processing: How focus and pitch accent determine the size of the N400 effect. *Neuropsychologia*, *49*(5), 813-820.

Wang, J., Wong, W. K., Wang, S., & Chen, H. C. (2017). Primary phonological planning units in spoken word production are language-specific: Evidence from an ERP study. *Scientific Reports*, *7*, 5815–5822.

Wong, A. W. K, & Chen, H.-C. (2008). Processing segmental and prosodic information in Cantonese word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1172–1190.

Wong, A. W. K, & Chen, H.-C. (2009). What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review*, *16*, 888–892.

Wong, A. W. K., Huang. J, & Chen, H.-C. (2012). Phonological units in spoken word production: Insights from Cantonese. *PloS One*, *7*(11):e48776. Doi: 10.1371/journal.pone.0048776.

Wong, A. W. K., Wang, J., Ng, T. Y. & Chen, H.-C. (2016). Syllabic encoding during overt speech production in Cantonese: Evidence from temporal brain responses. *Brain Research*, *1648*, 101–109.

You, W., Zhang Q., & Verdonschot, R. G. (2012). Masked syllable priming effects in word and picture naming in Chinese. *PLoS ONE*, *7*(10):e46595. Doi: 10.1371/journal.pone.0046595.

Yu, M., Mo, C. & Mo, L. (2014). The role of phoneme in Mandarin Chinese production: Evidence from ERPs. *PLoS One*, *9*(9):e106486. Doi:10.1371/journal.pone.0106486.

Yue, Y., & Zhang, Q.F. (2015). Syllable and segments effects in Mandarin Chinese spoken word production. *Acta Psychologica Sinica,*(3), 319-328.

[岳源, & 张清芳. (2015). 汉语口语产生中音节和音段的促进和抑制效应. *心理学报*(3), 319-328.]

Zhang, Q.F., & Yang, Y. (2005). The Phonological Planning Unit in Chinese Monosyllabic Word Production. *Psychological Science (China)*, *28*, 374–378.

[张清芳, & 杨玉芳. (2005). 汉语单音节词汇产生中音韵编码的单元. *心理科学*, *28*(2), 374-378.]

Zhang, Q. F. (2008). Phonological encoding in monosyllabic and bisyllabic Mandarin word production: Implicit priming paradigm study. *Acta Psychologica Sinica*, *40*(3), 253–262.

[张清芳. (2008). 汉语单音节和双音节词汇产生中的音韵编码过程: 内隐启动范式研究. 心理学报, *40*(3), 253–262.]

Zhang, Q., & Damian, M. F. (2019). Syllables constitute proximate units for Mandarin speakers: Electrophysiological evidence from a masked priming task. *Psychophysiology*, e13317. doi:10.1111/psyp.13317

Primary phonological planning units in Chinese speech word production: evidence from an ERP study with implicit priming paradigm

WANG Xuejiao, ZHANG Qingfang

(Department of Psychology, Renmin University of China, Beijing 100872, China)

Abstract

Languages may differ in terms of the proximate units of phonological encoding in spoken word production. It is widely accepted that phonemes are the primary units used in Indo-European languages. However, it is controversial what the functional units (syllables or phonemes) are in Chinese spoken word production. In the framework of *proximate unit principle* hypothesis for Mandarin Chinese spoken word production, phonological content is activated in a syllable while its syllabic frame is retrieved. Both syllabic content and frame are linked sequentially and tone is also specified at this point. Crucially, syllables are retrieved at an early stage of word form encoding, and then the phonemes of the syllable are retrieved in parallel and linked to their position in the syllable frame. The present study investigated the temporal courses of syllables and phonemes in word production using an implicit priming paradigm, and electrophysiological signals were measured concurrently. Specifically, participants were young native Chinese speakers with low level of English ability for excluding the influence of the second language.

In an implicit priming paradigm, four sets of word pairs served as experimental stimuli. Each set consisted of four pairs, and the first word of a pair was the cue word, while the second was the response word. In the homogeneous conditions, four response words in a set shared certain phonological properties. For instance, the four response words in a set were 危机(/wei1ji1/, crisis), 围墙 (/wei2qiang2/, bounding wall), 尾巴 (/wei3ba/, tail), and 卫星(/wei4xing1/, satellite), which share the first syllable but not tone. The heterogeneous conditions used the same word pairs, but assigned them to sets in which they didn't share any properties. Two homogeneous conditions (syllable shared and initial phoneme shared) were used in the study. Participants were asked to memorize the sets of four pairs of associate words, after which they were presented with the cue words, and were required to produce the response words.

Behavioral results didn't show any syllable or phoneme effects. EEG results revealed that syllable overlap conditions elicited significant effects in the time window of 100~400 ms, and the effect firstly showed up in the left and right anterior regions, following in the middle anterior region, and finally in the left and right anterior regions, right and middle posterior regions. While the phoneme effects only appeared between 500~600 ms after the cue words were presented, and were distributed in the right anterior and posterior regions. Both the syllabically and phonemically homogeneous conditions were characterized by more positive ERP waveforms when compared to the heterogeneous conditions. We suggested that syllable overlap effect arises at the early stage of phonological encoding, whereas the phoneme overlap effect possibly arises at the later stage of phonological encoding, or phonetic encoding in Chinese spoken word production. Our findings provide support for the assumption of the proximate unit principle.

Key words spoken word production; phonological encoding; proximate unit principle; syllable; phoneme

附录 1：实验中所使用的材料

			异质											
	组别	音节	1			2			3			4		
同质	/wei/		灾祸	危机	wei1	院子	围墙	wei2	屁股	尾巴	wei3	人造	卫星	wei4
	/hu/		医生	护士	hu4	氧气	呼吸	hu1	剃须	胡子	hu2	宝石	琥珀	hu3
	/guo/		收获	果实	guo3	惩罚	过错	guo4	烧水	锅炉	guo1	熊猫	国宝	guo2
	/fan/		神仙	凡人	fan2	化学	反应	fan3	食物	饭馆	fan4	航海	帆船	fan1
		音素	1			2			3			4		
	/t/		地下	通道	tong1	师傅	徒弟	tu2	运动	体育	ti3	宇宙	太空	tai4
	/l/		辩护	律师	lv4	废品	垃圾	la1	闪电	雷声	lei2	生日	礼物	li3
	/d/		披风	斗篷	dou3	和尚	道士	dao4	榔头	钉子	ding1	鸦片	毒品	du2
	/m/		排序	名次	ming2	帅哥	美女	mei3	视线	目光	mu4	擦拭	抹布	ma1